

Expression Identification and Emotional Classification of Students in Job Interviews Based on Image Processing



Lingling Cui¹, Weigong Kong², Yiming Sun¹, Lin Shao^{3*}

¹ Admission and Employment Department, Hebei Vocational University of Technology and Engineering, Xingtai 054000, China

² Economic Management Department, Hebei Vocational University of Technology and Engineering, Xingtai 054000, China

³ Department of Art and Media, Hebei Vocational University of Technology and Engineering, Xingtai 054000, China

Corresponding Author Email: shaolin7607@126.com

<https://doi.org/10.18280/ts.390228>

ABSTRACT

Received: 15 November 2021

Accepted: 8 February 2022

Keywords:

image processing, job interview, expression identification, emotional classification

The performance of college students in job interviews can be significantly promoted, if they are guided properly to identify and regulate negative emotions. However, the existing automatic expression identification algorithms cannot recognize expressions ideally, due to the small sample set, and the lack of diverse storage forms. To solve the problem, this paper explores the expression identification and emotional classification of students in job interviews based on image processing. Firstly, the ideas of interview emotion identification were expounded based on computer technology and image processing technology, and the college students' interview emotion regulation process was modeled. Then, the histogram of oriented gradients (HOG) was adopted to extract the local textures and edges from the expression images of students in job interviews, and the face expressions were identified for the analysis on interview emotions. Based on the graph neural network (GNN) and representation learning, a job interview expression identification algorithm was designed for college students, which effectively suppresses the uncertainty of these images in the real-world unconstrained environments. The proposed algorithm was proved effective through experiments.

1. INTRODUCTION

During the job hunt, college students are often anxious about interviews, due to the employment pressure and the lack of relevant experience [1-5]. Many studies have confirmed that those with a high interview anxiety perform poorly in job interviews, and receive relatively few offers [6-9]. The performance of college students in job interviews can be significantly promoted, if they are guided properly to identify and regulate negative emotion [10-16]. Many researchers have introduced face expression identification algorithms to the emotional analysis of people in various scenes [17-22]. These algorithms can classify the emotional states according to the degree of variation in expressions. There are three basic steps of these algorithms: image preprocessing, expression feature extraction, as well as expression identification and classification. The last step is of paramount importance.

Face expressions are crucial to the description of a person's emotions. The identification of face expressions has been widely applied in many fields, e.g., human-computer interaction, and driver state monitoring, and piqued the interest of researchers. After extracting face features, Wang et al. [23] selected features through the maximization of joint mutual information and conditional mutual information, maximization of correlation, and minimization of redundancy, and systematically compared the selected features. Multiple machine learning technologies were applied to classification. For the first time, they explored the gray level co-matrix and Haralick features of face expression recognition. Human

emotional recognition based on face expressions is of great significance in smart human-computer interaction applications. However, face images change greatly in the real environment, due to the complex background and brightness. To solve the problem, Choudhary and Shukla [24] proposed a robust face detection method based on skin tone enhancement model, and designed a face expression identification algorithm based on block principal component analysis (PCA). The brightness range of face images was expanded by the homomorphic filter, and the face expressions were recognized by the block PCA of deep neural network (DNN).

Huang et al. [25] presented an accurate, innovative, real-time emotional identification and recording system based on the novel deep learning technology. A nine-layer convolutional neural network (CNN) was established. The accuracy was improved by the addition of backpropagation. Besides, the front faces were detected by Haar cascades and opensource computer vision [26]. To solve the absence of emotions and improve teaching effect, Liang [27] devised an intelligent teaching method based on face expression recognition. The traditional active shape model was improved to extract feature points of faces, and the face expressions were recognized based on the geometric features of faces and the support vector machine (SVM). Bobkowska et al. [28] highlighted the importance of using high-speed cameras to record the development of face emotions: these cameras can detect micro-emotions. In addition, the face image was divided into eight feature segments. It was proved that the expression speed at each position of the face, and the maximum change

of specific emotions both vary with the moments of emotions.

Facing emotional analysis, face expression identification has largely developed through two stages: psychological research, and emotion recognition. In the early stage, the expression identification and emotional analysis rely on experts, which are time-consuming, laborious, and ineffective. Currently, automatic expression identification algorithms are in vogue. These algorithms are based on computers and image processing. However, few of them can recognize expressions ideally, due to the small sample set, and the lack of diverse storage forms. To accurately recognize students' expressions and classify their emotions in job interviews, this paper explores the expression identification and emotional classification of students in job interviews based on image processing. Section 2 explains the ideas of interview emotion identification based on computer technology and image processing technology, and models the college students' interview emotion regulation process. Section 3 uses the histogram of oriented gradients (HOG) to extract the local textures and edges from the expression images of students in job interviews, and identifies the face expressions for the analysis on interview emotions. Based on the graph neural network (GNN) and representation learning, Section 4 designs a job interview expression identification algorithm for college students, which effectively suppresses the uncertainty of these images in the real-world unconstrained environments. Finally, the effectiveness of our algorithm was demonstrated through experiments.

2. RESEARCH ROADMAP

With the development of emotional recognition practices, scholars in the field of computer technology and image processing have proposed various emotional analysis and identification methods based on different emotional theories. These methods provide an important technical foundation for emotional recognition in various application scenarios. Figure

1 explains the roadmap for interview emotion identification based on computer technology and image processing technology, which mainly includes verbal behavior, face expressions, body poses, pronunciations and tones, problem-solving ideas, and multi-modal hybrid data processing.

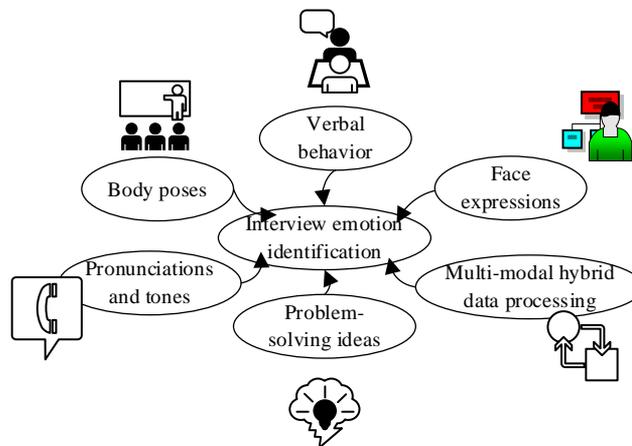


Figure 1. Roadmap for interview emotion identification based on computer technology and image processing technology

Situational interview, the most reliable interview method, is commonly adopted by job interviewers for college students. The interviewers raise a series of questions about the scenes related to the position or work. The answers to these questions are provided by an expert panel or job interviewers before the interview. During the interview, the college student should regulate his/her stressful emotions, as he/she listens to the questions (Figure 2). As he/she answers the questions, he/she needs to regulate emotional responses like psychological perception, behavioral expression, and physical response, and complete the interview with proper manners, accurate language expression, and clear answers.

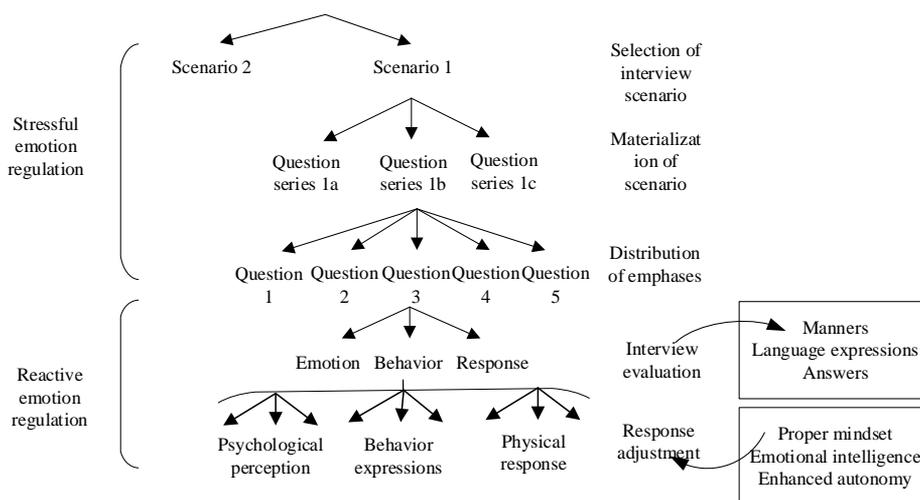


Figure 2. Emotional regulation process model for college students in interviews

3. FACE EXPRESSION IDENTIFICATION

The HOG algorithm can accurately depict the shape and contours of image targets. As a result, it has been widely applied to image retrieval and classification, as well as target

identification and tracking. This paper adopts the HOG to extract the local textures and edges of students' face expression images in job interviews, because important parts of the face change with face expressions.

The HOG feature extraction of students' interview

expression images is detailed as follows: The collected images may be too bright or too dark. Thus, the Gamma correction method is adopted to normalize the interview expression images. The overall brightness of the images is therefore adjusted. Let $HD(a, b)$ and $HD_0(a, b)$ be the gray value of a pixel in the original image and in the normalized image, respectively; α be the correction coefficient. Then, we have:

$$HD_0(a, b) = HD(a, b)^\alpha \quad (1)$$

To obtain the gradient size and gradient direction of an interview expression image, it is necessary to calculate the image gradients. The horizontal and vertical templates are $[-1, 0, 1]$ and $[-1, 0, 1]^T$, respectively. The horizontal gradient $TD_a(a, b)$ and vertical gradient $TD_b(a, b)$ of the interview expression image can be respectively calculated by:

$$TD_a(a, b) = HD_0(a+1, b) - HD_0(a-1, b) \quad (2)$$

$$TD_b(a, b) = HD_0(a, b+1) - HD_0(a, b-1) \quad (3)$$

The gradient size and gradient direction of the interview expression image can be respectively calculated by:

$$TD(a, b) = \sqrt{TD_a(a, b)^2 + TD_b(a, b)^2} \quad (4)$$

$$\beta(a, b) = \tan^{-1} \left(\frac{TD_b(a, b)}{TD_a(a, b)} \right) \quad (5)$$

After the calculation, the detection window is divided into a series of grids, and the HOG of each grid is solved. Then, adjacent grids are merged into a large block, and the HOG of each block is calculated. Finally, the HOGs of all blocks are serial connected to obtain the HOG feature histogram of the interview expression image. Figure 3 explains the process of acquiring the HOG feature histogram.

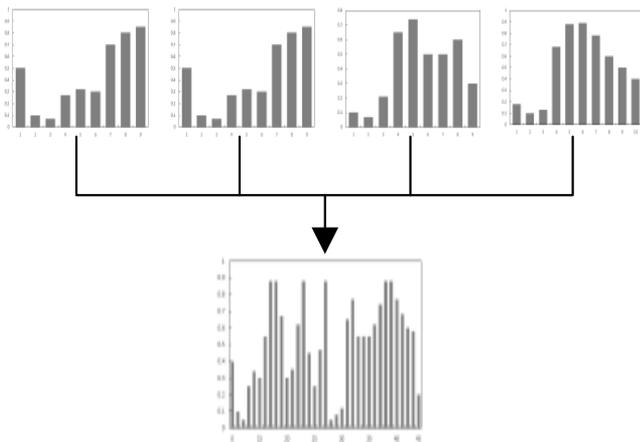


Figure 3. Acquisition process of the HOG feature histogram

The ability of a HOG feature to express the interview expression image depends on the scale. The HOG features of multiple scales should be fused to extract richer and more comprehensive features of students' expressions in job interviews. To shorten the computing time and improve expression recognition, the important information of interview

expression images is retained, while the redundant information of the fused multiscale HOG feature is removed. This paper adopts the PCA to reduce the dimensionality of the fused multiscale HOG feature. The specific steps of the algorithm are as follows:

Suppose there are N student's interview expression images, each of which contains M expression features $A=[a_1, a_2, \dots, a_N]^T$. The dimensionality of A is $N \times M$. The expression feature of the i -th image is denoted by a . Then, the mean of A can be calculated by:

$$\bar{a} = \frac{1}{N} \sum_{i=1}^N a_i \quad (6)$$

Then, each sample is subtracted by the mean:

$$a'_i = a_i - \bar{a} \quad (7)$$

Based on a'_i , a new feature set matrix $C=[a'_1, a'_2, \dots, a'_N]^T$ is established, whose dimensionality is $N \times M$. The covariance matrix $D=C^T C$ is solved, whose dimensionality is $M \times M$. After that, the eigenvalue μ and eigenvector w of D are solved, forming the dimensionally reduced matrix. Then, the top- S μ are ranked in descending order, and the w corresponding to μ are composed into a matrix $W=[w_1, w_2, \dots, w_S]$, whose dimensionality is $M \times S$. The S -th principal component is denoted as w_S . The dimensionally reduced feature set B can be obtained by multiplying A with W :

$$B = AW \quad (8)$$

The dimensionality of B is $N \times S$. Through the above steps, the dimensionality reduction is completed by the PCA.

4. MODELING AND OPTIMIZATION OF LABEL DISTRIBUTION

In the real-world unconstrained environment, the students' interview expression images have complex lighting conditions and head poses. The prior knowledge of the images captured in the controlled lab environment cannot meet the needs of expression identification in the real-world environment. This paper proposes an identification algorithm for students' interview expressions based on the GNN and representation learning, which effectively suppresses the uncertainty brought by the students' interview expression images in the real-world unconstrained environment.

Our GNN can describe the hidden connections between face action units and expressions of the students. The face action units are treated as the nodes of the GNN. The node representation is updated based on the information transmitted between the nodes. Let $V=(v_1, v_2, \dots, v_m) \in R^{m \times e(0)}$ be the features of face action units; m be the number of such features; e be the dimensionality of the features. Each layer of the GNN can be expressed as:

$$F^{(k+1)} = \rho(XF^{(k)}Q^{(k)}) \quad (9)$$

Let $X \in R^{m \times m}$ be the correlation matrix of face action units; $\rho(\cdot)$ be the activation function; $F^{(k)} \in R^{m \times e(k)}$ be the activation matrix of the k -th layer, with $F^{(0)}=V$; $Q^{(k)} \in R^{e(k) \times e(k+1)}$ be the

trainable weight matrix of the k-th layer; $F^{e(k+1)} \in R^{m \times e(k+1)}$ be the activated output of the k+1-th layer. The correlation between face action units can be expressed as a conditional probability $GL(XV_j|XV_i)$, i.e., the occurrence probability of XV_j in the appearance of XV_i . $GL(XV_j|XV_i)$ is not equal to $GL(XV_i|XV_j)$. To obtain matrix $N \in R^{m \times m}$ for building the correlation matrix, it is necessary to compute the number of occurrences of each XV. Let N_{ij} be the co-occurrence times of XV_i and XV_j ; M_i be the number of XV_i ; $GL_{ij}=GL(XV_j|XV_i)$ be the occurrence probability of XV_j in the appearance of XV_i . Then, the conditional probability can be calculated by:

$$GL_i = \frac{N_i}{M_i} \quad (10)$$

The correlation matrix X can be calculated by:

$$X_{ij} = \begin{cases} \frac{GL_{ij}}{\sum_{j=1}^m GL_{ij}}, i \neq j \\ 1 - GL_{ij}, i = j \end{cases} \quad (11)$$

For a given students' interview expression image A , the d classes of basic expressions can be expressed as $B=\{b_1, b_2, \dots, b_d\}$, and the distribution of emotional labels related to A_i can be expressed as $E_i=\{e_{i1}, e_{i2}, \dots, e_{id}\}$.

Let ω_{MX} be the model parameter; g_{MX} be the eigenvector passing through the convolutional layer; g_{HH} be the eigenvector passing through the mixed pooling layer. Under the fundamental framework of ResNet-50, this paper adds a mixed pooling layer consisting of max pooling and average pooling. The image level feature U can be obtained by this layer:

$$U = g_{HH} (g_{MX} (A; \omega_{MX})) \in R^E \quad (12)$$

According to the image representations involved in network training and the preset weight matrix, the probability distribution of students' interview emotions can be predicted by:

$$b_i^* = Q^T U \quad (13)$$

In addition, the interview emotion label $b_i \in R^d$ is given to the students' interview expression image A_i , where b_{ij} is the occurrence probability of each emotion. Then, the normalized distribution of interview expression labels can be expressed as:

$$h_{ij} = \frac{\exp(b_{ij}^*)}{\sum_l \exp(b_{lj}^*)} \quad (14)$$

Considering the confidence of the emotion labels in the sample set of the real-world unconstrained environment, the weight coefficient σ is defined to balance the distribution of interview emotion labels and the given labels. Then, the new distribution of interview emotion labels can be expressed as:

$$e_{ij} = \sigma g_{ij} + (1 - \sigma) b_{ij} \quad (15)$$

The distribution network of interview emotion labels is

trained by normal classification loss method:

$$K = -\frac{1}{d} \sum_j b_{ij} \ln \hat{b}_{ij} \quad (16)$$

After obtaining the distribution of students' interview emotion labels, the students' interview expression image A_i is provided with the distribution E_i of emotion labels. In our network, the activation value of the last layer, i.e., the fully connected layer, is $a=\Phi(A; \omega)$, where ω is the network parameter. Then, a can be converted into probability distribution GL by:

$$GL(b_j | A_j; \omega) = \frac{\exp(a_j)}{\sum_j \exp(a_j)} \quad (17)$$

For the training image set $P=\{(A_1, E_1), (A_2, E_2), \dots, (A_n, E_n)\}$, the network is trained to obtain weight ω . On this basis, a distribution E_i^* is generated to approximate the distribution E_i of students' interview emotion labels as much as possible. The distance between the two distributions can be quantified by Kullback-Leibler (KL) divergence:

$$KL(E_i || E_i^*) = \sum_i E_i \ln \frac{E_i}{E_i^*} \quad (18)$$

To prevent over-fitting, L2 regularization is introduced to our model. Let $\|\cdot\|_G$ be the norm. Then, we have:

$$K' = \frac{1}{2} \|E_i - E_i^*\|_G^2 \quad (19)$$

Let μ be the parameter of K2 regularization. Then, the optimal parameter ω_{best} can be defined as:

$$\omega_{best} = \underset{\omega}{\operatorname{argmin}} \sum_i E_i \ln \frac{E_i}{E_i^*} + \frac{\mu}{2} \sum_i \|E_i - E_i^*\|_G^2 \quad (20)$$

The overall loss of our model can be given by:

$$K(\omega) = -\sum_i \sum_j e_{ij} \ln GV(b_j | A_i, \omega) + \frac{\mu}{2} \left(\sum_i \sum_j e_{ij} - GV(b_j | A_i, \omega) \right)^2 \quad (21)$$

The loss can be minimized by the small batch gradient descent algorithm. Let β be the learning rate of the algorithm. Then, ω_j can be iteratively updated by:

$$\omega_j \leftarrow \omega_j - \beta \frac{\partial K(\omega)}{\partial \omega_j} \quad (22)$$

By the chain rule, the partial derivative of ω can be calculated by:

$$\frac{\psi K(\omega)}{\delta \omega_j} = -\sum_j \sum_j e_{ij} \frac{1}{GV(b_j | A_i, \omega)} \frac{\psi GV(b_j | A_i, \omega)}{\psi \omega_j} + \mu \sum_j \sum_j \left(e_{ij} - GV(b_j | A_i, \omega) \right) \frac{\psi GV(b_j | A_i, \omega)}{\psi \omega_j} \quad (23)$$

where,

$$\frac{\partial GV(b_j | A_i, \omega)}{\partial \omega_j} = GV(b_j | A_i, \omega) (\phi_{(j=i)} - GV(b_j | A_i, \omega)) \quad (24)$$

For any l and j , if the two are equal, then ϕ is 1; otherwise, ϕ is 0. Once the proper network parameter ω is obtained through network learning, the distribution E^* of emotion labels for any new sample A of students' interview expression image can be generated through the forward propagation in the model. Finally, the proposed model outputs the predicted distribution W_c^* of students' interview emotions, where:

$$c^* = \underset{i}{\operatorname{argmax}} \hat{E}_i \quad (25)$$

5. EXPERIMENTS AND RESULT ANALYSIS

This paper tests the fusion of multiscale HOG features. The HOG feature on each scale was dimensionally reduced through the PCA, before feature fusion. Figure 4 shows the optimal principal component contribution rate of dimensionally reduced HOG features, and the classification accuracy of students' interview emotions at different feature dimensions. The highest classification accuracy was defined as the accuracy corresponding to the optimal principal component contribution rate. The principal component contribution rate was preset to the range of (70, 100). The scale of HOG features was fixed as 10×10 . As shown in Figure 4, the classification accuracy of students' interview emotions

peaked at 86.64%, when the principal component contribution rate stood at around 80%.

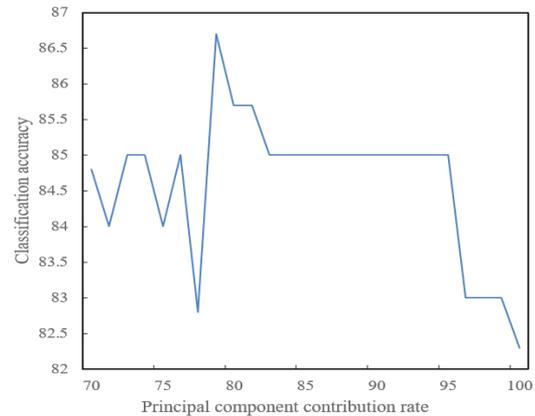


Figure 4. Principal component contribution rate of dimensionally reduced HOG features

Table 1 shows the experimental results on feature fusion at different scales. It can be seen that the students' interview emotions were best classified, after the features of two different scales, namely, 10×10 and 15×15 , were fused, and processed by the SVM.

To sum up, whether the features are fused with fixed or variable scales, the fused HOG feature can bring a higher classification accuracy of interview emotions than a single scale of HOG features, as long as the scales are appropriate. If the scales are not suitable, the interview emotions will not be classified correctly.

Table 1. Feature fusion results at different scales

Scale	Classification accuracy	Principal component contribution rate	Feature dimensionality
5+10	84.26%	81%	112
10+15	88.19%	83%	45
10+25	86.62%	86%	135
10+25+30	85.15%	88%	45
5+10+25+30	83.68%	82%	83

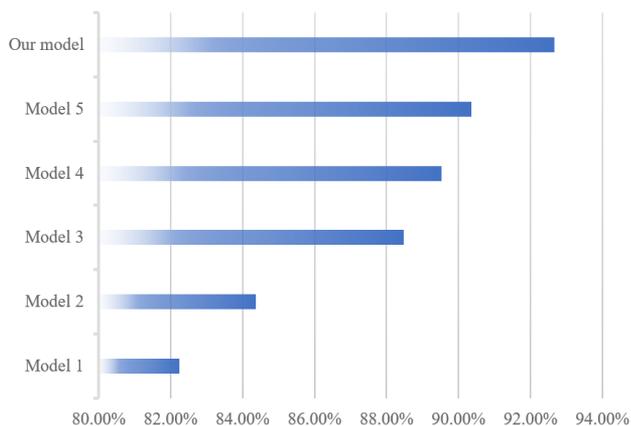


Figure 5. Interview emotion classification accuracies of different feature extraction methods

To further verify the effectiveness of our feature extraction algorithm, the classification accuracy of interview emotions of our algorithm was compared with that of histogram (Model 1),

local binary pattern (LBP) (Model 2), speeded-up robust features (SURF) (Model 3), scale-invariant feature transform (SIFT) (Model 4), and Gabor (Model 5). The comparison in Figure 5 shows that our feature extraction algorithm classified interview emotions better than other algorithms, and performed better in feature fusion.

Table 2. Mean classification accuracy of each algorithm

Contrastive algorithm	TSTNet	3DCNN	ConvLSTM	RAN
Emotion class	4 classes	4 classes	4 classes	4 classes
Cross validation	5 folds	5 folds	5 folds	5 folds
Accuracy (%)	83.62	85.19	83.05	82.69
Contrastive algorithm	SCN	MDMO	CNN+LSTM	Our model
Emotion class	4 classes	4 classes	4 classes	4 classes
Cross validation	5 folds	5 folds	5 folds	5 folds
Accuracy (%)	86.41	81.03	82.47	87.05

Furthermore, the proposed identification algorithm for students' interview expressions, which is based on GNN and representation learning, was compared with 7 common methods on a self-developed dataset. The contrastive methods include TSTNet, three-dimensional CNN (3D CNN), convolutional long short-term memory network (ConvLSTM), RAN, SCN, MDMO, and CNN+LSTM. As shown in Table 2, our algorithm accurately classified 87.05% of all interview emotions, higher than the accuracy of all contrastive methods. This means uncertainty can be effectively suppressed by setting up the distribution of emotion labels for highly uncertain samples. However, there are very few samples meeting the confidence standard for emotion labels, if the

confidence of emotion labels is considered for the sample set in the real-world unconstrained environment. In other words, although the face images with fuzzy emotions are correctly classified, over-fitting may occur if these images are used to train the model. This would hinder the model learning of highly discriminable face expression features.

Next, an analysis of variance with a single factor and multiple dependent variables was carried out, where the independent variable is the students' interview expression feature. Four classes of emotions were considered, including relaxed and positive, neutral, slightly nervous, and very anxious.

Table 3. Random test results of different groups

Variable		Relaxed and positive	Neutral	Slightly nervous	Very anxious
Group A N=22	Mean	42.16	25.61	15.92	36.58
	Standard deviation	9.48	7.49	4.75	6.37
Group B N=21	Mean	42.57	25.63	14.85	36.41
	Standard deviation	7.18	5.29	4.37	6.58
Group C N=21	Mean	45.69	23.84	13.86	33.62
	Standard deviation	0.92	0.97	0.61	0.58
Group D N=21	Mean	48.69	23.58	11.69	35.84
	Standard deviation	7.92	5.37	2.08	4.69
<i>F</i>		0.07	0.09	0.72	0.85
Significance		0.83	0.94	0.68	0.52

Table 4. Regulation effect of interview emotions in each group

Variable	Group A N=22		Group B N=21		Group C N=21		Group D N=21		<i>F</i>	Significance
	Mean	Standard deviation								
1	3.526	0.869	3.951	1.326	2.659	1.302	3.812	1.025	1.639	0.281
2	0.174	0.263	0.628	1.958	0.218	0.518	1.326	1.859	3.481	0.038
3	0.136	0.231	0.315	0.742	0.629	1.659	0.574	1.627	0.918	0.418
4	0.395	0.392	0.393	0.915	0.127	0.416	0.465	0.625	1.748	0.328
5	2.658	0.968	2.741	0.935	2.316	1.326	3.813	0.981	5.485	0.015
6	16.253	13.261	18.269	13.285	26.152	19.415	12.629	12.637	1.958	0.184
7	1.629	0.637	1.847	0.718	1.329	0.736	2.058	0.816	2.847	0.048
8	58.162	23.694	68.326	24.625	78.415	36.295	65.748	18.629	1.182	0.326
9	2.957	1.282	2.693	1.695	1.694	1.485	3.052	1.748	3.265	0.041

As shown in Table 3, the four groups of students from different majors did not have significant differences in the relaxed and positive emotion ($F=0.07, P>0.05$) and in the neutral emotion ($F=0.09, P>0.05$), and have very significant differences in the slightly nervous and very anxious emotions ($F=0.72, 0.85, P>0.05$). Therefore, the students from different majors differ in nervousness and anxiety during job interviews. The difference is largely dependent on the influence of different majors over the emotional management ability.

In addition, another analysis of variance with a single factor and multiple dependent variables was carried out, where the independent variable is the students' interview expression feature. The dependent variables are nine face expression index scores for the students' interview responses, including left and right eyebrows (variables 1 and 2), left and right brow ridges (variables 3 and 4), four sides of mouth corners (variables 5-8), and bottom of jaw (variable 9).

As shown in Table 4, concerning the interview responses, the four groups of students from different majors had very significant differences in variable 5 ($F=5.485, P<0.05$), significant differences in variables 2, 7, and 9 ($F=3.481, 2.847, 3.265, P<0.05$), and insignificant differences in other variables.

These results show that the students from different majors differ in the face expressions of nervousness and anxiety during job interviews. This agrees with the conclusions of the previous studies.

6. CONCLUSIONS

This paper explores the expression identification and emotional classification of students in job interviews based on image processing. After explaining the ideas of interview emotion identification based on computer technology and image processing technology, the authors modeled the college students' interview emotion regulation process. Then, the HOG was employed to extract the local textures and edges from the expression images of students in job interviews, and realized the identification of the face expressions for the analysis on interview emotions. In addition, a job interview expression identification algorithm was designed based on the GNN and representation learning. The algorithm can effectively control the uncertainty brought by students' interview expression images in the real-world unconstrained

environment.

In the experimental section, the authors firstly tested the fusion of multiscale HOG features, and summed up the classification accuracies of HOG features after dimensionality reduction. The results confirm that the fused HOG feature can bring a higher classification accuracy of interview emotions than a single scale of HOG features, as long as the scales are appropriate. Next, the classification accuracy of the proposed algorithm for interview emotions was compared with Histogram, LBP, SURF, SIFT, and Gabor. The comparison shows that that our feature extraction algorithm classified interview emotions better than other algorithms, and performed better in feature fusion. Afterwards, the proposed identification algorithm for students' interview expressions, which is based on GNN and representation learning, was compared with 7 common methods on a self-developed dataset. Our algorithm achieved the best classification performance on interview emotions. Finally, the authors performed analyses of variance with a single factor and multiple dependent variables. The results demonstrate the difference between students from different majors in nervousness and anxiety during job interviews.

REFERENCES

- [1] Jin, X., Bian, Y., Geng, W., Chen, Y., Chu, K., Hu, H., Yang, C. (2019). Developing an agent-based virtual interview training system for college students with high shyness level. In 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 998-999. <https://doi.org/10.1109/VR.2019.8797764>
- [2] Zhang, C.H., Yue, Y., Li, D.M. (2013). Research on social and environmental factors impact on college students' interview impression management behavior. In Proceedings of the International Conference on Information Engineering and Applications (IEA) 2012, pp. 11-17. https://doi.org/10.1007/978-1-4471-4847-0_2
- [3] Vilar, E., Noriega, P., Borges, T., Rebelo, F., Ramos, S. (2020). Can an environmental feature influence interview anxiety? In International Conference on Human-Computer Interaction, pp. 351-369. https://doi.org/10.1007/978-3-030-49757-6_25
- [4] Shimizu, S., Jincho, N., Kikuchi, H. (2019). Influence of interactive questions on the sense of presence and anxiety in a virtual-reality job-interview simulation. In Proceedings of the 2019 3rd International Conference on Virtual and Augmented Reality Simulations, pp. 1-5. <https://doi.org/10.1145/3332305.3332307>
- [5] Borges, T., Vilar, E., Noriega, P., Ramos, S., Rebelo, F. (2016). Virtual reality to study job interview anxiety: evaluation of virtual environments. In Advances in Ergonomics in Design, 25-33. https://doi.org/10.1007/978-3-319-41983-1_3
- [6] Andres, J.M., Hutt, S., Ocumpaugh, J., Baker, R.S., Nasiar, N., Porter, C. (2021). How anxiety affects affect: a quantitative ethnographic investigation using affect detectors and data-targeted interviews. In International Conference on Quantitative Ethnography, pp. 268-283. https://doi.org/10.1007/978-3-030-93859-8_18
- [7] Kwon, J.H., Powell, J., Chalmers, A. (2013). How level of realism influences anxiety in virtual reality environments for a job interview. International Journal of Human-Computer Studies, 71(10): 978-987. <https://doi.org/10.1016/j.ijhcs.2013.07.003>
- [8] Hall Jr, P., Gosha, K. (2018). The effects of anxiety and preparation on performance in technical interviews for HBCU computer science majors. In Proceedings of the 2018 ACM SIGMIS Conference on Computers and People Research, pp. 64-69. <https://doi.org/10.1145/3209626.3209707>
- [9] Hartholt, A., Mozgai, S., Rizzo, A.S. (2019). Virtual job interviewing practice for high-anxiety populations. In Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents, pp. 238-240. <https://doi.org/10.1145/3308532.3329417>
- [10] Aldeneh, Z., Jaiswal, M., Picheny, M., McInnis, M., Provost, E.M. (2019). Identifying mood episodes using dialogue features from clinical interviews. arXiv preprint arXiv:1910.05115. <https://doi.org/10.48550/arXiv.1910.05115>
- [11] Zhang, S., Bian, C., Mai, X., Qin, S., Liu, C. (2021). Regulation of overnight sleep on negative emotion memory with moral-violation. Kexue Tongbao, 66(13): 1642-1652. <https://doi.org/10.1360/TB-2020-0814>
- [12] Xie, S., Shi, M., Yan, H. (2018). A study of negative emotion regulation of college students by social games design. In International Conference on E-Learning and Games, pp. 313-317. https://doi.org/10.1007/978-3-030-23712-7_43
- [13] Price, M.J., Mudrick, N.V., Taub, M., Azevedo, R. (2018). The role of negative emotions and emotion regulation on self-regulated learning with MetaTutor. In International Conference on Intelligent Tutoring Systems, pp. 170-179. https://doi.org/10.1007/978-3-319-91464-0_17
- [14] Yin, N., Zhang, J.H., Wang, H.L., Wang, A.X., Xu, G.Z. (2021). Electroencephalogram source localization and brain network of magnetic stimulation at acupoints to regulate negative emotion. Transactions of China Electrotechnical Society, 36(4): 756-764. <https://doi.org/10.19595/j.cnki.1000-6753.tces.201256>
- [15] Bouazza, H., Bendella, F. (2017). Adaptation of a model of emotion regulation to modulate the negative emotions based on persistency. Multiagent and Grid Systems, 13(1): 19-30. <https://doi.org/10.3233/MGS-170259>
- [16] Zhang, W., Wang, M., Zhu, Y.C. (2020). Does government information release really matter in regulating contagion-evolution of negative emotion during public emergencies? From the perspective of cognitive big data analytics. International Journal of Information Management, 50: 498-514. <https://doi.org/10.1016/j.ijinfomgt.2019.04.001>
- [17] Buhari, A.M., Ooi, C.P., Baskaran, V.M., Phan, R.C., Wong, K., Tan, W.H. (2022). Invisible emotion magnification algorithm (IEMA) for real-time micro-expression recognition with graph-based features. Multimedia Tools and Applications, 81: 9151-9176. <https://doi.org/10.1007/s11042-021-11625-1>
- [18] Liu, C., Yang, J., Zhao, W.N., Zhang, Y.N., Shi, C.P., Miao, F.J., Zhang, J.S. (2021). Differential privacy protection of face images based on region growing. Traitement du Signal, 38(5): 1385-1401. <https://doi.org/10.18280/ts.380514>
- [19] Minaee, S., Minaei, M., Abdolrashidi, A. (2021). Deep-emotion: Facial expression recognition using attentional

- convolutional network. *Sensors*, 21(9): 3046. <https://doi.org/10.3390/s21093046>
- [20] Liu, D., Wang, Z., Wang, L., Chen, L. (2021). Multi-modal fusion emotion recognition method of speech expression based on deep learning. *Frontiers in Neurobotics*, 15: 697634. <https://doi.org/10.3389/fnbot.2021.697634>
- [21] Yu, J.Y., Bai, X.J. (2021). Analysis of classroom learning behaviors based on internet of things and image processing. *Traitement du Signal*, 38(3): 845-851. <https://doi.org/10.18280/ts.380331>
- [22] Zhao, Y., Chen, D. (2021). Expression EEG multimodal emotion recognition method based on the bidirectional LSTM and attention mechanism. *Computational and Mathematical Methods in Medicine*, 2021: Article ID 9967592. <https://doi.org/10.1155/2021/9967592>
- [23] Wang, X., Chen, X., Cao, C. (2020). Human emotion recognition by optimally fusing facial expression and speech feature. *Signal Processing: Image Communication*, 84: 115831. <https://doi.org/10.1016/j.image.2020.115831>
- [24] Choudhary, D., Shukla, J. (2020). Feature extraction and feature selection for emotion recognition using facial expression. In 2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM), pp. 125-133. <https://doi.org/10.1109/BigMM50055.2020.00027>
- [25] Huang, L., Xie, F., Zhao, J., Shen, S., Guang, W., Lu, R. (2020). Human emotion recognition based on face and facial expression detection using deep belief network under complicated backgrounds. *International Journal of Pattern Recognition and Artificial Intelligence*, 34(14): 2056010. <https://doi.org/10.1142/S0218001420560108>
- [26] Wei, Z. (2019). A novel facial expression recognition method for identifying and recording emotion. In IOP Conference Series: Materials Science and Engineering 612(5): 052048. <https://doi.org/10.1088/1757-899X/612/5/052048>
- [27] Liang, Y. (2019). Intelligent emotion evaluation method of classroom teaching based on expression recognition. *International Journal of Emerging Technologies in Learning*, 14(4): 127-141.
- [28] Bobkowska, K., Przyborski, M., Skorupka, D. (2018). Emotion recognition-the need for a complete analysis of the phenomenon of expression formation. In E3S Web of Conferences, 26: 00013. <https://doi.org/10.1051/e3sconf/20182600013>